

DENSITY OF SAMPLING AND INTERPOLATION IN REPRODUCING KERNEL HILBERT SPACES

HARTMUT FÜHR, KARLHEINZ GRÖCHENIG, ANTTI HAIMI, ANDREAS KLOTZ,
AND JOSÉ LUIS ROMERO

ABSTRACT. We derive necessary density conditions for sampling and for interpolation in general reproducing kernel Hilbert spaces satisfying some natural conditions on the geometry of the space and the reproducing kernel. If the volume of shells is small compared to the volume of balls (weak annular decay property) and if the kernel possesses some off-diagonal decay or even some weaker form of localization, then there exists a critical density D with the following property: a set of sampling has density $\geq D$, whereas a set of interpolation has density $\leq D$. The main theorem unifies many known density theorems in signal processing, complex analysis, and harmonic analysis. For the special case of bandlimited function we recover Landau's fundamental density result. In complex analysis we rederive a critical density for generalized Fock spaces. In harmonic analysis we obtain the first general result about the density of coherent frames.

1. INTRODUCTION

How many samples of a function f are necessary to completely recover f ? The first answer is the sampling theorem of Whittaker, Kotelnikov, Shannon, and others [49]. It provides an explicit and elegant reconstruction formula for the recovery of a bandlimited function from its samples on a grid and establishes a fundamental relation between the bandwidth of f and the sampling density (the Nyquist rate in engineering terminology). This sampling theorem is the basis of modern information theory [49] and remains the model for analog-digital and digital-analog conversion.

The decisive mathematical theorems work for more general notions of bandwidth and for non-uniform sampling and are due to Beurling [7, 44] (sufficient conditions) and Landau [30]. Landau's necessary conditions give a precise meaning to the concept of a Nyquist rate for bandlimited functions.

2010 *Mathematics Subject Classification.* 42C15, 94A12, 46C05, 42C30, 32A70.

Key words and phrases. Reproducing kernel Hilbert space, Beurling density, frame, sampling theorem, interpolation, square-integrable representation, Fock space, bandlimited function.

K. G. and A. K. were supported in part by the project P26273 - N25 of the Austrian Science Fund (FWF). J. L. R. gratefully acknowledges support from a Marie Curie fellowship, under grant PIIF-GA-2012-327063 (EU FP7). A. H. was supported by the Lise Meitner grant M1821 of the Austrian Science Fund (FWF).

To this day, Landau's theorem is the prototype of a density theorem, it has inspired several hundred papers on sampling. Landau's necessary conditions have been transferred, modified, and adapted to dozens of similar situations. Here is a short, but by no means exhaustive list of density theorems in the wake of Landau:

- (i) Sampling in spaces of analytic functions, in particular, in Bargmann-Fock space [32, 42, 45] and in generalized Fock spaces [1, 31, 35].
- (ii) Sampling of bandlimited functions with derivatives [24].
- (iii) Necessary density conditions of Gabor frames [12, 40]. This topic alone has attracted about hundred papers, for a detailed history of this density theorem we refer to Heil's survey article [25].
- (iv) Density conditions for abstract frames with some localization properties [4–6].
- (v) Sampling in spaces of bandlimited functions on Lie groups [19].
- (vi) Sampling in spaces of variable bandwidth [23].
- (vii) Sampling in spaces of bandlimited functions associated to an integral transform, e.g., the Hankel transform [2].
- (viii) Density of frames in the orbit of an irreducible unitary representation of a homogeneous nilpotent Lie group [26].

Essentially each of these contributions on necessary density conditions for sampling and interpolation uses and modifies one of three methods.

- (i) Landau's original method is based on the spectral analysis of a family of localization operators (composition of the projection onto bandlimited functions with a time-limiting operator). This method is very powerful, but can become quite technical. Usually, the generalization of Landau's method is difficult.
- (ii) The method of Ramanathan and Steger [40] was originally developed to prove the density theorem for Gabor frames. It compares and estimates the dimension of finite-dimensional subspaces corresponding to a local patch of a sampling set with the dimension of finite-dimensional subspaces corresponding to a local patch of an interpolating set. This is the method used most frequently. However, it is not universally applicable, because it requires the existence of a set that is simultaneously sampling and interpolating (or at least the construction of interpolating sets and sampling sets with almost the same density).
- (iii) A third method goes back to Kolountzakis and Lagarias [28] (proof of Lemma 2.3) who studied the density of tilings by translation. This method

was then used by Iosevich and Kolountzakis [27] to prove a version of Landau's theorem and by Nitzan-Olevski [34] to give the simplest proof of Landau's theorem. With hindsight this method consists of comparing a set of sampling to a continuous frame. For us this is the method of choice that we will use in our investigation.

We observe that all density theorems listed above treat certain Hilbert spaces with a reproducing kernel (or isomorphic copies thereof). This fact and the similarity of all proofs raises the question of a universal density theorem in reproducing kernel Hilbert spaces. This point of view leads immediately to the pertinent questions: what are the concrete conditions on the underlying configuration space and on the reproducing kernel to lead to a density theorem? What is the relevant density concept in a reproducing kernel Hilbert space? Is there a critical density in a reproducing kernel Hilbert space that separates sets of sampling from sets of interpolation?

In this paper we attempt to give an answer to these questions and will prove a general density theorem for functions in a reproducing kernel Hilbert space. Here is a simplified version of our main result.

Theorem 1.1. *Let X be a metric measure space with a metric d and a measure μ such that balls have finite measure, μ is non-degenerate, and satisfies the weak annular decay property, i.e., $\inf_x \mu(B_r(x)) > 0$ for some $r > 0$, and*

$$\limsup_{r \rightarrow \infty} \sup_{x \in X} \frac{\mu(B_r(x) \setminus B_{r-1}(x))}{\mu(B_r(x))} = 0.$$

Furthermore, let $\mathcal{H} \subseteq L^2(X, \mu)$ be a reproducing kernel Hilbert space with a reproducing kernel $k(x, y)$ satisfying $\inf_{x \in X} k(x, x) > 0$ and an off-diagonal decay condition of the form

$$(1) \quad |k(x, y)| \leq C(1 + d(x, y))^{-\sigma} \quad \text{for all } x, y \in X$$

for some $\sigma > 0$ satisfying $\lim_{r \rightarrow \infty} \sup_{x \in X} \int_{X \setminus B_r(x)} (1 + d(x, y))^{-2\sigma} d\mu(y) = 0$.

(i) Necessary conditions for sampling: If for $\Lambda \subset X$ there exist $A, B > 0$ such that

$$(2) \quad A\|f\|^2 \leq \sum_{\lambda \in \Lambda} |f(\lambda)|^2 \leq B\|f\|^2 \quad \text{for all } f \in \mathcal{H},$$

then

$$D^-(\Lambda) := \liminf_{r \rightarrow \infty} \inf_{x \in X} \frac{\#(\Lambda \cap B_r(x))}{\mu(B_r(x))} \geq \liminf_{r \rightarrow \infty} \inf_{x \in X} \frac{1}{\mu(B_r(x))} \int_{B_r(x)} k(y, y) d\mu(y).$$

(ii) Necessary conditions for interpolation: Likewise, let $\Lambda \subset X$ and assume that for every $a \in \ell^2(\Lambda)$, there exists a function $f \in \mathcal{H}$ such that

$$f(\lambda) = a_\lambda, \quad \lambda \in \Lambda,$$

then

$$D^+(\Lambda) := \limsup_{r \rightarrow \infty} \sup_{x \in X} \frac{\#(\Lambda \cap B_r(x))}{\mu(B_r(x))} \leq \limsup_{r \rightarrow \infty} \sup_{x \in X} \frac{1}{\mu(B_r(x))} \int_{B_r(x)} k(y, y) d\mu(y).$$

Following established terminology, we call a set $\Lambda \subseteq X$ that satisfies the sampling inequality (2) a set of (stable) sampling, while a set satisfying the interpolation property in part (ii) of Theorem 1.1 is called a set of interpolation. Alternatively, Λ is a set of sampling, if and only if $\{k_\lambda : \lambda \in \Lambda\}$ is a frame for \mathcal{H} , and Λ is a set of interpolation if and only if $\{k_\lambda : \lambda \in \Lambda\}$ is a Riesz sequence in \mathcal{H} , i.e., if there exist $A, B > 0$ such that

$$(3) \quad A\|c\|^2 \leq \left\| \sum_{\lambda \in \Lambda} c_\lambda k(\cdot, \lambda) \right\|_2^2 \leq B\|c\|^2 \quad \text{for all } c \in \ell^2(\Lambda).$$

The densities $D^-(\Lambda)$ and $D^+(\Lambda)$ are the obvious generalizations of the lower and upper Beurling density to metric spaces.

The principal merit of Theorem 1.1 is the clarification of the main notions that go into a density theorem. To prove a density theorem, one needs

- (i) geometric data and the compatibility of metric and measure, and
- (ii) estimates for the reproducing kernel.

The verification of these properties is by no means trivial. Indeed, kernel estimates (Bergman, Bargmann, and other reproducing kernels) constitute a deep and rich area of analysis. Theorem 1.1 shifts the emphasis in proofs of density theorems: it is important to understand the geometry and the reproducing kernel, but it is no longer necessary to prove a “new” density theorem from scratch with tedious modifications of known techniques.

As an example we show how Landau’s original theorem follows from Theorem 1.1. The discussion also shows some of the difficulties in applying Theorem 1.1.

Let $X = \mathbb{R}^d$ with Lebesgue measure μ and $\Omega \subseteq \mathbb{R}^d$ be a set of finite Lebesgue measure and define $B_\Omega = \{f \in L^2(\mathbb{R}^d) : \text{supp } \hat{f} \subseteq \Omega\}$ to be the subspace of band-limited functions with spectrum in Ω . Then

$$f(x) = \int_{\Omega} \hat{f}(\xi) e^{2\pi i x \cdot \xi} d\xi = \int_{\mathbb{R}^d} f(y) \int_{\Omega} e^{2\pi i \xi(x-y)} d\xi dy,$$

and therefore B_Ω is a reproducing kernel Hilbert space with reproducing kernel $k(x, y) = \int_{\Omega} e^{2\pi i \xi(x-y)} d\xi = \widehat{1_\Omega}(y - x)$. Clearly \mathbb{R}^d satisfies all geometric conditions of Theorem 1.1. The computation of the averaged trace is equally easy, since $k(x, x) = \int_{\Omega} 1 d\xi = \mu(\Omega)$ and thus $\frac{1}{\mu(B_r(x))} \int_{B_r(x)} k(y, y) d\mu(y) = \mu(\Omega)$ independent of x and r . Thus a set of sampling must satisfy $D^-(\Lambda) \geq \mu(\Omega)$ and a set of interpolation $D^+(\Lambda) \leq \mu(\Omega)$, which is Landau’s theorem. The decay condition (1) is satisfied for simple spectra, e.g., when Ω is a cube or a convex set with smooth boundary. Yet we must be cautious: in general the kernel does not satisfy the decay condition in (1), because the Fourier transform of the characteristic function of a

compact sets may decay arbitrarily slowly [20]. In the main theorem (Theorem 2.2) we will impose a much weaker condition on the kernel and thus will take care of this subtle point.

Most density theorems of the list above can be understood as an example of the general density theorem in reproducing kernel Hilbert spaces. To demonstrate the wide applicability of Theorem 1.1 we will rederive some of the fundamental density theorems in several areas of analysis.

(i) *Signal analysis*: as already indicated, Theorem 1.1 implies Landau’s necessary density conditions for bandlimited functions.

(ii) *Complex analysis* (in several variables): we will deduce Lindholm’s density conditions [31] for generalized Fock spaces.

(iii) *Harmonic analysis*: We will derive a necessary condition for the density of a frame in the orbit of a square-integrable, unitary representation of a group of polynomial growth. A special case of this result is the density theorem for Gabor frames.

Theorem 1.1 is definitely not the end of density theorems. Of course, it is our main ambition to prove new density results. In this sense the axiomatic set-up serves as a preparation for future work. Currently there are numerous results about sampling theorems for “sufficiently dense” sets. See for instance, [16, 36–39]. These results need to be complemented by a critical density, provided that it exists at all.

Finally, let us point out some limitations. We mention that the weak annular decay property of the measure is not always satisfied, as it is tied to the growth of balls in X . Thus Theorem 1.1 excludes a number of very interesting examples, for instance, density theorems in Bergman spaces [41, 43, 44] and the density of wavelet frames [29]. However, in these cases the Beurling density is not the correct density, and to this date it is an open problem whether a critical density always exists in this context. In our view the Beurling density is the correct notion of density in geometries with polynomial (or subexponential) growth, whereas in geometries with exponential growth new phenomena arise (which are not at all understood).

The paper is organized as follows: In Section 2 we collect the set of assumptions for a general density theorem and then formulate the main result. Section 3 contains a discussion of the main hypotheses and several technical lemmas. In Section 4 we provide the proof of the main density theorem. In Section 5 we apply the density theorem to rederive several fundamental density theorems from the literature.

2. THE DENSITY THEOREM

2.1. Assumptions. We first list the general assumptions on the geometry and the reproducing kernel to make a density theorem work.

(A) **Assumptions on the metric and the measure.** We assume that (X, d) is a metric space, and μ is a measure on X with the following properties:

- The metric $d : X \times X \rightarrow [0, +\infty)$ is a $\mu \otimes \mu$ measurable function and the balls $B_r(x) = \{y \in X : d(y, x) < r\}$ satisfy $\mu(B_r(x)) < \infty$ for all $r > 0$ and $x \in X$.
- *Non-degeneracy of balls* (Axiom NDB): There exist $r > 0$ such that

$$\inf_{x \in X} \mu(B_r(x)) > 0.$$

- *Weak annular decay property* (Axiom WAD): Spherical shells have small volume compared to full balls:

$$(4) \quad \limsup_{r \rightarrow \infty} \sup_{x \in X} \frac{\mu(B_{r+1}(x) \setminus B_r(x))}{\mu(B_r(x))} = 0.$$

(B) **Assumptions on the reproducing kernel.** We assume that $\mathcal{H} \subseteq L^2(X, \mu)$ is a reproducing kernel Hilbert space with reproducing kernel $k(x, y)$ so that

$$f(x) = \int_X k(x, y) f(y) d\mu(y) = \langle f, k_x \rangle,$$

where $k_x(y) = k(y, x) = \overline{k(x, y)}$. The assumptions on the kernel are as follows:

- Condition on the diagonal (Axiom D): there exist constants $C_1, C_2 > 0$ such that for all $x \in X$

$$C_1 \leq k(x, x) \leq C_2.$$

- *Weak localization of the kernel* (Axiom WL): For every $\epsilon > 0$ there is a constant $r = r(\epsilon)$, such that

$$(5) \quad \sup_{x \in X} \int_{X \setminus B_r(x)} |k(x, y)|^2 d\mu(y) < \epsilon^2.$$

- *The homogeneous approximation property* (Axiom HAP): Assume that Λ is such that $\{k_\lambda : \lambda \in \Lambda\}$ is a Bessel sequence for \mathcal{H} , i.e., $\sum_{\lambda \in \Lambda} |f(\lambda)|^2 \leq C \|f\|^2$ for all $f \in \mathcal{H}$. Then for every $\epsilon > 0$ there is a constant $r = r(\epsilon)$, such that

$$(6) \quad \sup_{x \in X} \sum_{\lambda \in \Lambda \setminus B_r(x)} |k(x, \lambda)|^2 < \epsilon^2.$$

We note that this version of axiom (HAP) differs from the usual homogeneous approximation property used in the literature. Compare the discussion in [4].

2.2. Sets, densities and traces. A set $\Lambda \subseteq X$ is called *relatively separated* if there exists $\rho_0 > 0$ such that for all $\rho \geq \rho_0$, there exists $C = C_\rho > 0$ such that:

$$(7) \quad \#(\Lambda \cap B_\rho(x)) \leq C_\rho \mu(B_\rho(x)), \quad x \in X.$$

Definition 2.1. The lower Beurling density of a set $\Lambda \subseteq X$ is defined to be

$$(8) \quad D^-(\Lambda) = \liminf_{r \rightarrow \infty} \inf_{x \in X} \frac{\#(\Lambda \cap B_r(x))}{\mu(B_r(x))},$$

and the upper Beurling density of Λ is

$$(9) \quad D^+(\Lambda) = \limsup_{r \rightarrow \infty} \sup_{x \in X} \frac{\#(\Lambda \cap B_r(x))}{\mu(B_r(x))}.$$

Note that for relatively separated sequences, the upper Beurling density is finite by the non-degeneracy of the balls.

We will compare the density of a set of sampling or interpolation to an invariant of the reproducing kernel Hilbert space. The correct Nyquist rate is the averaged trace of the kernel. We define the lower and upper traces as follows:

$$(10) \quad \text{tr}^-(k) = \liminf_{r \rightarrow \infty} \inf_{x \in X} \frac{1}{\mu(B_r(x))} \int_{B_r(x)} k(y, y) d\mu(y),$$

$$(11) \quad \text{tr}^+(k) = \limsup_{r \rightarrow \infty} \sup_{x \in X} \frac{1}{\mu(B_r(x))} \int_{B_r(x)} k(y, y) d\mu(y).$$

2.3. Main result. In this general context one can prove the following necessary density conditions for sets of sampling and interpolation.

Theorem 2.2. Assume that \mathcal{H} is a reproducing kernel Hilbert space of functions on a metric measure space X satisfying the assumptions of Section 2.1.

(i) If $\Lambda \subseteq X$ is a set of stable sampling for \mathcal{H} , then

$$(12) \quad D^-(\Lambda) \geq \text{tr}^-(k) \text{ and } D^+(\Lambda) \geq \text{tr}^+(k).$$

(ii) If $\Lambda \subseteq X$ is a set of interpolation for \mathcal{H} , then

$$(13) \quad D^-(\Lambda) \leq \text{tr}^-(k) \text{ and } D^+(\Lambda) \leq \text{tr}^+(k).$$

It is instructive to work out the trivial case of a space X with finite diameter, i.e., $X = B_R(x)$ for all $x \in X$. Assuming the existence of a set Λ of sampling in this case, we observe by Lemma 3.6 below that Λ is finite, hence \mathcal{H} is finite-dimensional. Choosing any orthonormal basis $\varphi_1, \dots, \varphi_m$ of \mathcal{H} , we have

$$k(x, y) = \sum_{j=1}^m \varphi_j(x) \overline{\varphi_j(y)}.$$

In particular, we get

$$\text{tr}^-(k) = \text{tr}^+(k) = \frac{1}{\mu(X)} \int_X k(y, y) d\mu(y) = \frac{m}{\mu(X)}.$$

On the other hand, given any subset Λ , we have

$$D^+(\Lambda) = D^-(\Lambda) = \frac{\#\Lambda}{\mu(X)}.$$

Hence Theorem 2.2 says that $\#\Lambda \leq m$ is a necessary condition for sets of interpolation, whereas $\#\Lambda \geq m$ is necessary for sets of sampling (=uniqueness). Of course, both results follow from an elementary dimension count: The cardinality of a Riesz sequence is bounded by the dimension of the underlying vector space, whereas a frame must contain a basis and its cardinality exceeds the dimension of \mathcal{H} .

3. DISCUSSION OF THE ASSUMPTIONS AND PRELIMINARY LEMMAS

In the following we always assume the axioms from Section 2.1. We discuss the axioms and prove some easy consequences.

3.1. Metric and measure. Most of the geometric conditions are technical conditions to exclude pathologies.

First note that we *do not* assume that the σ -algebra of μ -measurable sets - the domain of μ - is generated by the class of open sets associated with the metric d . Rather, we only assume that the function d is $\mu \otimes \mu$ measurable. This means that the sets $\{(x, y) \in X \times X : d(x, y) < r\}$ belong to the smallest σ -algebra generated by the sets $A \times B$, with A, B in the domain of μ . In particular, the balls $B_r(x)$ are μ measurable. However, more general sets that are open with respect to the metric d may not be μ measurable, since they may fail to be a countable union of balls.

The decisive condition is the weak annular decay property (WAD). This condition links the metric and the measure and imposes some compatibility between them. Even in simple examples, axiom (WAD) requires some care. For example, the standard metric $d_1(x, y) = |x - y|$ on \mathbb{R} with Lebesgue measure fulfills the weak annular decay property, but the topologically equivalent metric $d_2(x, y) = \log(1 + |x - y|)$ violates (WAD).

For complicated geometries the verification of the weak annular decay property is decidedly non-trivial. In the literature one often uses the stronger *annular decay property* [46]. A metric measure space (X, d, μ) satisfies this property, if there exist constants $C > 0$ and $\delta \in (0, 1]$ such that for every $h \in [0, 1]$, $r > 0$, $x \in X$

$$(14) \quad \mu(B_r(x) \setminus B_{(1-h)r}(x)) \leq Ch^\delta \mu(\overline{B_r}(x)).$$

The annular decay property implies the weak annular decay property. Indeed, if (14) holds, the choice $h = r^{-1}$ shows that $\mu(B_r(x) \setminus B_{r-1}(x)) \leq Cr^{-\delta} \mu(\overline{B_r}(x))$. In addition, for $r \geq 1$ and $\varepsilon > 0$,

$$\begin{aligned} \mu(\overline{B_r}(x)) &\leq \mu(B_{r+\varepsilon}(x)) \leq \mu(B_r(x)) + C\varepsilon^\delta (r + \varepsilon)^{-\delta} \mu(\overline{B_r}(x)) \\ &\leq \mu(B_r(x)) + C\varepsilon^\delta \mu(\overline{B_r}(x)). \end{aligned}$$

So, choosing ε small enough ($\varepsilon = (2C)^{-1/\delta}$), we see that $\mu(\overline{B_r}(x)) \leq C\mu(B_r(x))$ and the conclusion follows.

In the literature one finds several conditions that imply the annular decay property, for instance, if X is a length space or if X has monotone geodesics. See [9, 46] for further discussion and more information on the annular decay property. Axiom (WAD) seems to be tied to the growth of balls and seems compatible with at most subexponential growth.

We now note that the measure μ is *locally doubling at large scales*.

Lemma 3.1. *There exist $r_0 > 0$ such that for all $r \geq r_0$, there is a constant $C_r > 0$ such that*

$$(15) \quad \mu(B_{2r}(x)) \leq C_r \mu(B_r(x)) \quad \text{for all } x \in X.$$

Proof. By Axiom (WAD), there exist r_0 , such that for $r \geq r_0$ and all $x \in X$, $\mu(B_{r+1}(x) \setminus B_r(x)) \leq \mu(B_r(x))$. As a consequence, $\mu(B_{r+1}(x)) \leq 2\mu(B_r(x))$. Iterating this estimate we conclude that $\mu(B_{2r}(x)) \leq C_r \mu(B_r(x))$, where $C_r := 2^{\lceil r \rceil}$. ■

Remark 3.2. Note that the proof of Lemma 3.1 only depends on Axiom (WAD).

The locally doubling property in (15) is much weaker than the usual doubling property for measures. See [11, 47] for more on locally doubling spaces.

We next formulate two lemmas on the number of points in and the measure of general “spherical” shells.

Lemma 3.3. *Let $\Lambda \subseteq X$ be relatively separated. Then for all sufficiently large $\rho > 0$, and $R > \rho$, $r > 0$, $x \in X$, we have*

$$(16) \quad \# \left(\Lambda \cap (B_{R+r}(x) \setminus B_R(x)) \right) \leq C_{\rho, \Lambda} \mu(B_{R+r+\rho}(x) \setminus B_{R-\rho}(x)),$$

where the constant $C_{\rho, \Lambda}$ depends only on ρ and Λ .

Proof. By Lemma 3.1, for sufficiently large ρ , the locally doubling property

$$(17) \quad \mu(B_\rho(x)) \leq C_{\rho/2} \mu(B_{\rho/2}(x))$$

and the estimate in (7) hold.

Let $\{B_{\rho/2}(y) : y \in X_0\}$ be a maximal packing of $B_{R+r}(x) \setminus B_R(x)$. This means that (i) $X_0 \subseteq B_{R+r}(x) \setminus B_R(x)$, (ii) $\{B_{\rho/2}(y) : y \in X_0\}$ is a disjoint family of balls and (iii) the family is maximal with respect to the properties (i) and (ii). By maximality, $\{B_\rho(y) : y \in X_0\}$ is a covering of $B_{R+r}(x) \setminus B_R(x)$. Using (7) and (17), we obtain

$$\begin{aligned} \#(\Lambda \cap (B_{R+r}(x) \setminus B_R(x))) &\leq \# \left(\bigcup_{y \in X_0} \Lambda \cap B_\rho(y) \right) \leq C \sum_{y \in X_0} \mu(B_\rho(y)) \\ &\leq CC_{\rho/2} \sum_{y \in X_0} \mu(B_{\rho/2}(y)) = CC_{\rho/2} \mu \left(\bigcup_{y \in X_0} B_{\rho/2}(y) \right) \\ &\leq CC_{\rho/2} \mu(B_{R+r+\rho}(x) \setminus B_{R-\rho}(x)), \end{aligned}$$

where C provided by (7) depends only on ρ and Λ . ■

We will apply the weak annular decay property in the following versions.

Lemma 3.4. *If (X, d, μ) satisfies the weak annular decay property, then, for all $\rho' > 0$,*

$$(18) \quad \limsup_{r \rightarrow \infty} \sup_{x \in X} \frac{\mu(B_{r+\rho'}(x))}{\mu(B_r(x))} = 1 ,$$

and

$$(19) \quad \limsup_{r \rightarrow \infty} \sup_{x \in X} \frac{\mu(B_{r+\rho'}(x) \setminus B_{r-\rho'}(x))}{\mu(B_r(x))} = 0 .$$

Proof. For the proof of (18), we observe that weak annular decay is equivalent to

$$\limsup_{r \rightarrow \infty} \sup_{x \in X} \frac{\mu(B_{r+1}(x))}{\mu(B_r(x))} = 1 .$$

For given $\rho' > 0$, we set $N = \lceil \rho' \rceil$ and obtain

$$\begin{aligned} 1 &\leq \sup_{x \in X} \frac{\mu(B_{r+\rho'}(x))}{\mu(B_r(x))} \leq \sup_{x \in X} \frac{\mu(B_{r+N}(x))}{\mu(B_r(x))} \\ &\leq \prod_{j=1}^N \sup_{x \in X} \underbrace{\frac{\mu(B_{r+j}(x))}{\mu(B_{r+j-1}(x))}}_{\rightarrow 1, \text{ as } r \rightarrow \infty} . \end{aligned}$$

This proves (18). For the proof of (19) note that

$$\begin{aligned} 0 &\leq \sup_{x \in X} \frac{\mu(B_{r+\rho'}(x) \setminus B_{r-\rho'}(x))}{\mu(B_r(x))} \\ &\leq \sup_{x \in X} \frac{\mu(B_{r+\rho'}(x))}{\mu(B_r(x))} - \inf_{x \in X} \frac{\mu(B_{r-\rho'}(x))}{\mu(B_r(x))} \end{aligned}$$

which converges to $1 - 1 = 0$ by (18), as $r \rightarrow \infty$. ■

3.2. The reproducing kernel. In practice, the upper bound in Axiom (D) and the weak localization property follow from off-diagonal decay estimates for the kernel (see Section 5). To verify the lower bound in Axiom (D), the following observation can be useful.

Lemma 3.5. *Let k be a reproducing kernel on $X \times X$. If there is a constant $C > 0$ such that for all $x \in X$, there exists $f_x \in X$ such that $\|f_x\| \leq C$ and $f_x(x) = 1$, then the lower bound in Axiom (D) holds.*

Proof. The claim follows from $1 = f_x(x) = \langle f_x, k_x \rangle \leq C \|k_x\| = C k(x, x)^{1/2}$. ■

Normalization of the reproducing kernel. In some examples, e.g., in Fock spaces of entire functions, the reproducing kernel is unbounded. This situation can be dealt with by the following normalization. Let $\psi : X \rightarrow \mathbb{R}^+$ and define a new measure $\tilde{\mu} = \psi^2 \mu$. Then J defined by $Jf = f\psi^{-1}$ is an isometry from $L^2(X, \mu)$ onto $L^2(X, \tilde{\mu})$ and $\tilde{\mathcal{H}} = J\mathcal{H} \subseteq L^2(X, \tilde{\mu})$ is again a reproducing kernel Hilbert space. We calculate the new reproducing kernel \tilde{k} as follows:

$$Jf(x) = \psi(x)^{-1}f(x) = \psi(x)^{-1}\langle f, k_x \rangle = \psi(x)^{-1}\langle Jf, Jk_x \rangle,$$

whence the new reproducing kernel is

$$(20) \quad \tilde{k}(x, y) = \psi(x)^{-1}\psi(y)^{-1}k(x, y) \quad x, y \in X.$$

In particular, if we choose $\psi(x) = \|k_x\|$, then $\tilde{k}(x, y) = \|k_x\|^{-1}\|k_y\|^{-1}k(x, y)$ and thus $\tilde{k}(x, x) = 1$. Axiom (D) can therefore always be fulfilled by a renormalization of the kernel. However, this may be at the price of destroying some other required properties. Note that in this normalization the critical density is always one, provided that Theorem 2.2 is still applicable.

Lemma 3.6. *Let $\Lambda \subseteq X$ be a set such that $\{k_\lambda : \lambda \in \Lambda\}$ is a Bessel sequence in \mathcal{H} . Then Λ is relatively separated.*

Proof. Let $\epsilon^2 := \frac{1}{2} \inf_{x \in X} k(x, x)$. By Axiom (D), $\epsilon > 0$. Select $r_1 = r_1(\epsilon)$ according to Axiom (WL). Since $\int_X |k(x, y)|^2 d\mu(y) = \|k_x\|^2 = k(x, x)$, we obtain that, for $\rho \geq r_1$,

$$(21) \quad \int_{B_\rho(x)} |k(x, y)|^2 d\mu(y) = k(x, x) - \int_{X \setminus B_\rho(x)} |k(x, y)|^2 d\mu(y) \geq \epsilon^2.$$

By hypothesis, there exists a constant $C > 0$ such that for all $f \in \mathcal{H}$

$$\sum_{\lambda \in \Lambda} |f(\lambda)|^2 \leq C \|f\|^2,$$

and this holds in particular for $f = k_y$. Using Lemma 3.1, select r_0 such that (15) holds for $r \geq r_0$. Let $x \in X$ and $\rho \geq \rho_0 := \max\{r_0, r_1\}$. Using (21) and Axiom (D), we estimate

$$\begin{aligned} \epsilon^2 \#(\Lambda \cap B_\rho(x)) &\leq \sum_{\lambda \in \Lambda \cap B_\rho(x)} \int_{B_\rho(\lambda)} |k(\lambda, y)|^2 d\mu(y) \\ &\leq \sum_{\lambda \in \Lambda \cap B_\rho(x)} \int_{B_{2\rho}(x)} |k(\lambda, y)|^2 d\mu(y) \leq \int_{B_{2\rho}(x)} \sum_{\lambda \in \Lambda} |k_y(\lambda)|^2 d\mu(y) \\ &\leq C \sup_{y \in X} \|k_y\|^2 \mu(B_{2\rho}(x)) \leq CC_\rho \mu(B_\rho(x)). \end{aligned}$$

Hence, (7) holds. ■

Remark 3.7. Note that the proof of Lemma 3.6 does not use Axiom (HAP).

We will need the following elementary facts about frames and Riesz sequences in a reproducing kernel Hilbert space \mathcal{H} . They are copied from [23, 34].

Lemma 3.8. *The following properties hold.*

(i) *Assume that $\{k_\lambda : \lambda \in \Lambda\}$ is a frame for \mathcal{H} with canonical dual frame $\{g_\lambda : \lambda \in \Lambda\}$. Then k_λ and g_λ satisfy the following:*

$$(22) \quad \sum_{\lambda \in \Lambda} k_\lambda(y) \overline{g_\lambda(y)} = k(y, y), \quad y \in X,$$

$$(23) \quad \sup_{y \in X} \sum_{\lambda \in \Lambda} |g_\lambda(y)|^2 < \infty \quad y \in X,$$

$$(24) \quad \sup_{\lambda \in \Lambda} \|g_\lambda\| = C < \infty,$$

$$(25) \quad \sup_{\lambda \in \Lambda} |\langle k_\lambda, g_\lambda \rangle| \leq 1.$$

(ii) *If $\{k_\lambda : \lambda \in \Lambda\}$ is a Riesz basis for a subspace $V \subseteq \mathcal{H}$ with biorthogonal basis $\{g_\lambda : \lambda \in \Lambda\} \subseteq V$, then (23), (24) hold true, while (22) is replaced by the inequality*

$$(26) \quad \sum_{\lambda \in \Lambda} k_\lambda(y) \overline{g_\lambda(y)} \leq k(y, y),$$

and in (25) holds the equality

$$(27) \quad \langle g_\lambda, k_\lambda \rangle = 1 \text{ for all } \lambda \in \Lambda.$$

Proof. The proof from [23, 34] is included for completeness. Let P_V be the orthogonal projection on the subspace V of \mathcal{H} . Identity (26) follows from

$$\begin{aligned} \sum_{\lambda \in \Lambda} k_\lambda(y) \overline{g_\lambda(y)} &= \sum_{\lambda \in \Lambda} \langle k_\lambda, k_y \rangle \langle k_y, g_\lambda \rangle \\ &= \left\langle \sum_{\lambda \in \Lambda} \langle k_y, g_\lambda \rangle k_\lambda, k_y \right\rangle = \langle P_V k_y, k_y \rangle \\ &\leq \|k_y\|^2 = k(y, y). \end{aligned}$$

The proof of (22) is the same, except that $P_V = I$ for frames and thus equality holds in the last step.

Item (23) follows from

$$\sum_{\lambda \in \Lambda} |g_\lambda(y)|^2 = \sum_{\lambda \in \Lambda} |\langle g_\lambda, k_y \rangle|^2 \leq C \|k_y\|^2 = C k(y, y),$$

where C is the upper frame bound for $\{g_\lambda : \lambda \in \Lambda\}$, and $k(y, y)$ is uniformly bounded by Axiom (D).

Item (25) is an immediate consequence of the minimality of the ℓ^2 -norms of the coefficients in the canonical frame expansion [13]:

$$k_{\lambda'} = \sum_{\lambda \in \Lambda} \langle k_{\lambda'}, g_{\lambda} \rangle k_{\lambda} = 1 \cdot k_{\lambda'} \quad \text{for every } \lambda' \in \Lambda,$$

so

$$|\langle k_{\lambda'}, g_{\lambda'} \rangle|^2 \leq \sum_{\lambda \in \Lambda} |\langle k_{\lambda'}, g_{\lambda} \rangle|^2 \leq 1 \quad \text{for every } \lambda' \in \Lambda.$$

Finally (24) is a general fact about frames. ■

4. PROOF OF THEOREM 2.2

In this section we prove the necessary density conditions for sets of sampling or interpolation in general reproducing kernel Hilbert spaces satisfying the conditions of Section 2.1. Similar to [27, 34], our proof is inspired by the method of Kolountzakis and Lagarias [28]. It is modeled on our own version in [23].

Proof. Step 1. The first part of the proof works both for sets of sampling and for sets of interpolation. Equivalently, we assume that $\{k_{\lambda} : \lambda \in \Lambda\}$ is either a frame with *canonical dual frame* $\{g_{\lambda} : \lambda \in \Lambda\}$ or that $\{k_{\lambda} : \lambda \in \Lambda\}$ is a Riesz sequence for some subspace $V \subseteq \mathcal{H}$ with *biorthogonal basis* $\{g_{\lambda} : \lambda \in \Lambda\} \subseteq V$. By Lemma 3.6, the set Λ is relatively separated. Let $\rho > 0$ be a suitably large radius, such that the conclusion of Lemma 3.3 holds.

In both cases we estimate the quantity

$$(28) \quad \int_{B_r(x)} \sum_{\lambda \in \Lambda} k_{\lambda}(y) \overline{g_{\lambda}(y)} d\mu(y)$$

for large r .

Fix $\epsilon > 0$ and $x \in X$, and choose $R = R(\epsilon)$ such that both kernel axioms WL (5) and HAP (6) are satisfied. In the proof, we will just write B_r for the ball $B_r(x)$ to abbreviate the notation.

We partition Λ and write accordingly

$$\begin{aligned} \sum_{\lambda \in \Lambda} k_{\lambda}(y) \overline{g_{\lambda}(y)} &= \left(\sum_{\lambda \in \Lambda \cap B_{r-R}} + \sum_{\lambda \in \Lambda \cap (X \setminus B_{r+R})} + \sum_{\lambda \in \Lambda \cap (B_{r+R} \setminus B_{r-R})} \right) k_{\lambda}(y) \overline{g_{\lambda}(y)} \\ &= A_1(y) + A_2(y) + A_3(y). \end{aligned}$$

Estimate of $\int_{B_r} A_2$. Note that $y \in B_r$ and $\lambda \in \Lambda \setminus B_{r+R}$ implies that $d(\lambda, y) > R$. Then Axiom HAP (6) ensures that $\sum_{\lambda \in \Lambda \setminus B_{r+R}} |k(y, \lambda)|^2 \leq \sum_{\lambda \in \Lambda \setminus B_R(y)} |k(y, \lambda)|^2 < \epsilon^2$. Consequently, using also (23), we obtain

$$(29) \quad \left| \int_{B_r} A_2(y) d\mu(y) \right| \leq \int_{B_r} \left(\sum_{\lambda \in \Lambda \cap (X \setminus B_{r+R})} |k_{\lambda}(y)|^2 \right)^{1/2} \left(\sum_{\lambda \in \Lambda} |g_{\lambda}(y)|^2 \right)^{1/2} d\mu(y) \leq \epsilon C_2 \mu(B_r).$$

Estimate of $\int_{B_r} A_3$. For the third term observe that

$$(30) \quad \begin{aligned} \int_{B_r} |A_3(y)| d\mu(y) &\leq \sum_{\lambda \in \Lambda \cap (B_{r+R} \setminus B_{r-R})} \int_X |k_\lambda(y)| |g_\lambda(y)| d\mu(y) \\ &\leq \sum_{\lambda \in \Lambda \cap (B_{r+R} \setminus B_{r-R})} \|k_\lambda\| \|g_\lambda\|. \end{aligned}$$

Using Axiom (D) and the boundedness of the canonical dual frame (24), we obtain

$$(31) \quad \int_{B_r} |A_3(y)| d\mu(y) \leq C_3 \#(\Lambda \cap (B_{r+R} \setminus B_{r-R})).$$

Estimate of $\int_{B_r} A_1$. Next we estimate $|\int_{B_r} A_1(y) d\mu(y)|$. We write

$$\int_{B_r} A_1(y) d\mu(y) = \int_X \sum_{\lambda \in \Lambda \cap B_{r-R}} k_\lambda(y) \overline{g_\lambda(y)} d\mu(y) - \int_{X \setminus B_r} \sum_{\lambda \in \Lambda \cap B_{r-R}} k_\lambda(y) \overline{g_\lambda(y)} d\mu(y),$$

and set

$$L = \sum_{\lambda \in \Lambda \cap B_{r-R}} \int_{X \setminus B_r} k_\lambda(y) \overline{g_\lambda(y)} d\mu(y).$$

Then

$$(32) \quad \int_{B_r} A_1(y) d\mu(y) = \sum_{\lambda \in \Lambda \cap B_{r-R}} \langle k_\lambda, g_\lambda \rangle - L.$$

If $\lambda \in \Lambda \cap B_{r-R}$ and $y \in X \setminus B_r$, then $d(\lambda, y) > R$. Therefore the kernel axiom WL (6) and (24) imply that a single term contributing to L is majorized by

$$(33) \quad \left| \int_{X \setminus B_r} k_\lambda(y) \overline{g_\lambda(y)} d\mu(y) \right| \leq \left(\int_{X \setminus B_R(\lambda)} |k_\lambda(y)|^2 d\mu(y) \right)^{1/2} \|g_\lambda\| \leq \epsilon C''.$$

This estimate implies

$$(34) \quad |L| \leq \epsilon C_1 \#(\Lambda \cap B_{r-R}) \leq \epsilon C_1 \#(\Lambda \cap B_r).$$

From now on we distinguish the case of sets of sampling from sets of interpolation.

Step 2. Assume first that $\{k_\lambda : \lambda \in \Lambda\}$ is a Riesz sequence in \mathcal{H} . We rewrite the expansion

$$\int_{B_r} \sum_{\lambda \in \Lambda} k_\lambda(y) \overline{g_\lambda(y)} d\mu(y) = \int_{B_r} \sum_{j=1}^3 A_j(y) d\mu(y),$$

and, with the help of (32), (26), and (27), we obtain the estimate

$$\begin{aligned}
\#(\Lambda \cap B_{r-R}) &= \sum_{\lambda \in \Lambda \cap B_{r-R}} \langle k_\lambda, g_\lambda \rangle \\
&= \int_{B_r} A_1(y) d\mu(y) + L \\
&= \int_{B_r} \sum_{\lambda \in \Lambda} k_\lambda(y) \overline{g_\lambda(y)} d\mu(y) - \int_{B_r} A_2(y) d\mu(y) - \int_{B_r} A_3(y) d\mu(y) + L \\
&\leq \int_{B_r} k(y, y) d\mu(y) + \left| \int_{B_r} A_2(y) d\mu(y) \right| + \left| \int_{B_r} A_3(y) d\mu(y) \right| + |L|.
\end{aligned}$$

Using $\#(\Lambda \cap B_r) = \#(\Lambda \cap B_{r-R}) + \#(\Lambda \cap (B_r \setminus B_{r-R}))$ and the estimates for $\int_{B_r} A_j(y) d\mu(y)$ (see (29), (31) and (34)), we obtain that

$$\begin{aligned}
(35) \quad \#(\Lambda \cap B_r) &\leq \int_{B_r} k(y, y) d\mu(y) + \epsilon C_2 \mu(B_r) + C_3 \#(\Lambda \cap (B_{r+R} \setminus B_{r-R})) \\
&\quad + \epsilon C_1 \#(\Lambda \cap B_r) + \#(\Lambda \cap (B_r \setminus B_{r-R})).
\end{aligned}$$

Lemma 3.3 bounds the last term by

$$\#(\Lambda \cap (B_r \setminus B_{r-R})) \leq \#(\Lambda \cap (B_{r+R} \setminus B_{r-R})) \leq C_{\rho, \Lambda} \mu(B_{r+R+\rho} \setminus B_{r-R-\rho}),$$

so we conclude that

$$(36) \quad (1 - \epsilon C_1) \frac{\#(\Lambda \cap B_r)}{\mu(B_r)} \leq \frac{1}{\mu(B_r)} \int_{B_r} k(y, y) d\mu(y) + \epsilon C_2 + (1 + C_3) C_{\rho, \Lambda} \frac{\mu(B_{r+R+\rho} \setminus B_{r-R-\rho})}{\mu(B_r)}.$$

We recall that $B_r = B_r(x)$, take the supremum over all $x \in X$, let r tend to ∞ , and use Lemma 3.4 to deduce

$$\begin{aligned}
(1 - \epsilon C_1) D^+(\Lambda) &\leq \text{tr}^+(k) + \epsilon C_2 + (1 + C_3) C_{\rho, \Lambda} \limsup_{r \rightarrow \infty} \sup_{x \in X} \frac{\mu(B_{r+R+\rho}(x) \setminus B_{r-R-\rho}(x))}{\mu(B_r(x))} \\
&= \text{tr}^+(k) + \epsilon C_2.
\end{aligned}$$

Since $\epsilon > 0$ is arbitrary, it follows that $D^+(\Lambda) \leq \text{tr}^+(k)$ for every interpolating set Λ . The inequality $D^-(\Lambda) \leq \text{tr}^-(k)$ follows from (36) in a similar way, just taking inf instead of sup and lim inf instead of lim sup.

Step 3. Assume next that $\{k_\lambda : \lambda \in \Lambda\}$ is a frame for \mathcal{H} . Then by Lemma 3.8, (22) and (25), we have

$$\sum_{\lambda \in \Lambda} k_\lambda(y) \overline{g_\lambda(y)} = k(y, y)$$

and

$$|\langle k_\lambda, g_\lambda \rangle| \leq 1.$$

Proceeding as in Step 2 we obtain with $x \in X$ fixed and $B_r = B_r(x)$ that

$$\begin{aligned}
\int_{B_r} k(y, y) d\mu(y) &= \int_{B_r} \sum_{\lambda \in \Lambda} k_\lambda(y) \overline{g_\lambda(y)} d\mu(y) = \int_{B_r} \sum_{j=1}^3 A_j(y) d\mu(y) \\
&= \sum_{\lambda \in \Lambda \cap B_{r-R}} \langle k_\lambda, g_\lambda \rangle - L + \int_{B_r} A_2(y) d\mu(y) + \int_{B_r} A_3(y) d\mu(y) \\
&\leq \#(\Lambda \cap B_{r-R}) + \epsilon C_1 \#(\Lambda \cap B_r) + \epsilon C_2 \mu(B_r) + C_3 \#(\Lambda \cap (B_{r+R} \setminus B_{r-R})) \\
&\leq (1 + \epsilon C_1) \#(\Lambda \cap B_r) + \epsilon C_2 \mu(B_r) + C_3 C_{\rho, \Lambda} \mu(B_{r+R+\rho} \setminus B_{r-R-\rho}).
\end{aligned}$$

Consequently

$$(37) \quad \frac{1}{\mu(B_r)} \int_{B_r} k(y, y) d\mu(y) \leq (1 + \epsilon C_1) \frac{\#(\Lambda \cap B_r)}{\mu(B_r)} + \epsilon C_2 + C_3 C_{\rho, \Lambda} \frac{\mu(B_{r+R+\rho} \setminus B_{r-R-\rho})}{\mu(B_r)}.$$

Again, we take the infimum over all $x \in X$ and let r tend to ∞ to obtain via Lemmas 3.3 and 3.4

$$\mathrm{tr}^-(k) \leq (1 + \epsilon C_1) D^-(\Lambda) + \epsilon C_2.$$

Since $\epsilon > 0$ was arbitrary, the necessary density is $D^-(\Lambda) \geq \mathrm{tr}^-(k)$, as claimed. As in Step 2, the statement involving the upper trace and density follows by just taking sup instead of inf and lim sup instead of lim inf. \blacksquare

Remark 4.1. Instead of the Beurling densities one may also apply an ultra-filter to (36) and (37) and obtain a density theorem with respect to a so-called frame measure function. See [5, 6] for the notion of frame measure function and its applications to the comparison of frames.

Remark 4.2. Theorem 2.2 (ii) is valid without axiom (HAP).

Our proof of the density theorem emphasized the symmetry between sampling and interpolation. We have seen that the *same* estimates are used in both density theorems. If we give up this symmetry, we can streamline the proof of the interpolation part a bit, and deduce the density conditions without assuming the kernel axiom (HAP).

Indeed, with the notation of the preceding proof let $V_{B_r} = \mathrm{span}\{k_\lambda : \lambda \in \Lambda \cap B_r\}$ and P the orthogonal projection onto V_{B_r} . The (unique) biorthogonal basis in V_{B_r} is $\{Pg_\mu : g_\mu \in \Lambda \cap B_r\}$, because $\langle k_\lambda, Pg_\mu \rangle = \langle k_\lambda, g_\mu \rangle = \delta_{\lambda, \mu}$ for $\lambda, \mu \in \Lambda \cap B_r$. Choose $R = R(\epsilon)$ so that that axiom (WL) is satisfied. Then

$$\#(\Lambda \cap B_r) = \sum_{\lambda \in \Lambda \cap B_r} \langle k_\lambda, Pg_\lambda \rangle = \left(\int_{B_{r+R}} + \int_{X \setminus B_{r+R}} \right) \left(\sum_{\lambda \in \Lambda \cap B_r} k_\lambda(y) \overline{Pg_\lambda(y)} \right) d\mu(y) = I + \tilde{L}.$$

By (33) and (34) we obtain $|\tilde{L}| \leq \epsilon C_1 \#(\Lambda \cap B_r)$, whereas (26) yields

$$I = \int_{B_{r+R}} k(y, y) d\mu(y) \leq \int_{B_r} k(y, y) d\mu(y) + \sup_{x \in X} k(y, y) \mu(B_{r+R} \setminus B_r).$$

Consequently,

$$\#(\Lambda \cap B_r) \leq \int_{B_r} k(y, y) d\mu(y) + C_0 \mu(B_{r+R} \setminus B_r) + \epsilon C_1 \#(\Lambda \cap B_r),$$

which readily yields

$$D^\pm(\Lambda) \leq \text{tr}^\pm(k).$$

4.1. Off-diagonal decay with respect to a metric. In applications, the reproducing kernel often possesses some off-diagonal decay. In this case the kernel axioms are easier to check. The following proposition shows that Theorem 1.1 is a special case of Theorem 2.2.

Proof of Theorem 1.1. We show that the hypothesis of Section 2.1 are satisfied. The assumption on d and k clearly implies the weak localization condition (WL) and the diagonal condition (D). It only remains to check the homogeneous approximation property (HAP). Assume that $\Lambda \subseteq X$ is such that $\{k_\lambda : \lambda \in \Lambda\}$ is a Bessel sequence.

As noted in Remark 3.7, the proof of Lemma 3.6 does not depend on Axiom (HAP). Hence we can invoke that lemma to obtain $\rho > 0$ such that (7) holds. Similarly, we can invoke Lemma 3.1 - which, as noted in Remark 3.2, only depends on Axiom (WAD) - to further grant that

$$(38) \quad \mu(B_{2\rho}(x)) \leq C_\rho \mu(B_\rho(x)), \quad x \in X.$$

We observe first that the obvious inequality

$$1 + d(x, y) \leq (1 + d(x, \lambda))(1 + d(\lambda, y)) \quad \text{for all } x, y, \lambda \in X$$

implies that

$$(1 + d(x, \lambda))^{-2\sigma} \leq (1 + d(\lambda, y))^{2\sigma} (1 + d(x, y))^{-2\sigma}.$$

Therefore,

$$\begin{aligned} |k(x, \lambda)|^2 &\leq C(1 + d(x, \lambda))^{-2\sigma} = \frac{C}{\mu(B_\rho(\lambda))} \int_{B_\rho(\lambda)} (1 + d(x, \lambda))^{-2\sigma} d\mu(y) \\ &\leq \frac{C}{\mu(B_\rho(\lambda))} \int_{B_\rho(\lambda)} \frac{(1 + d(\lambda, y))^{2\sigma}}{(1 + d(x, y))^{2\sigma}} d\mu(y) \\ &\leq C \frac{(1 + \rho)^{2\sigma}}{\mu(B_\rho(\lambda))} \int_{B_\rho(\lambda)} (1 + d(x, y))^{-2\sigma} d\mu(y). \end{aligned}$$

Consequently,

$$\begin{aligned} & \sum_{\lambda \in \Lambda \cap (X \setminus B_r(x))} |k(x, \lambda)|^2 \\ & \leq C(1 + \rho)^{2\sigma} \int_X \left(\sum_{\lambda \in \Lambda \setminus B_r(x)} \mu(B_\rho(\lambda))^{-1} \mathbf{1}_{B_\rho(\lambda)}(y) \right) (1 + d(x, y))^{-2\sigma} d\mu(y). \end{aligned}$$

We note that the sum vanishes if $d(x, y) \leq r - \rho$, thus the integral can be taken over the set $X \setminus B_{r-\rho}(x)$. Next we estimate the sum for fixed $y \in X \setminus B_{r-\rho}(x)$. Note that if $y \in B_\rho(\lambda)$, then $B_\rho(y) \subseteq B_{2\rho}(\lambda)$, and, by (38),

$$\mu(B_\rho(\lambda)) \geq C_\rho^{-1} \mu(B_{2\rho}(\lambda)) \geq C_\rho^{-1} \mu(B_\rho(y)).$$

Hence, using (7), we can estimate

$$\sum_{\lambda \in \Lambda \setminus B_r(x)} \frac{1}{\mu(B_\rho(\lambda))} \mathbf{1}_{B_\rho(\lambda)}(y) \leq \frac{C_\rho}{\mu(B_\rho(y))} \sum_{\lambda \in \Lambda} \mathbf{1}_{B_\rho(y)}(\lambda) \leq CC_\rho.$$

In conclusion,

$$\sum_{\lambda \in \Lambda \setminus B_r(x)} |k(x, \lambda)|^2 \leq CC_\rho(1 + \rho)^{2\sigma} \int_{X \setminus B_{r-\rho}(x)} (1 + d(x, y))^{-2\sigma} d\mu(y).$$

By hypothesis, this expression tends to zero uniformly in x as $r \rightarrow \infty$, whence k satisfies (HAP). ■

5. EXAMPLES

In this section we discuss several examples of density theorems from different areas of analysis. Our point is to show that some of the fundamental density theorems in signal analysis, complex analysis, frame theory, and harmonic analysis follow from the axiomatic approach. All we have to do is to check the general conditions of Section 2.1 and formulate the corresponding theorem. This is not always easy, and our discussion will point out some of the difficulties and pitfalls.

5.1. Bandlimited Functions. Let $\Omega \subseteq \mathbb{R}^d$ be measurable with finite Lebesgue measure and $B_\Omega = \{f \in L^2(\mathbb{R}^d) : \text{supp } \hat{f} \subseteq \Omega\}$ be the corresponding Paley-Wiener space. As observed in the introduction, its reproducing kernel is

$$k(x, y) = \int_{\Omega} e^{-2\pi i \xi \cdot (y-x)} d\xi.$$

Clearly $X = \mathbb{R}^d$ with the Euclidean distance and Lebesgue measure $d\mu(x) = dx$ satisfy the geometrical assumptions. As for the kernel, we have $k(x, x) = |\Omega|$ and thus the averaged trace is obviously $\text{tr}^+(k) = \text{tr}^-(k) = \mu(\Omega)$.

The verification of the weak localization of the kernel is easy, because $k(x, y) = \widehat{1_\Omega}(y - x)$ where $\widehat{1_\Omega}$ is the Fourier transform of an L^2 -function. Therefore

$$\int_{\mathbb{R}^d \setminus B_r(x)} |\widehat{1_\Omega}(y - x)|^2 dy = \int_{|y| \geq r} |\widehat{1_\Omega}(y)|^2 dy < \epsilon^2$$

for suitably large r . The axiom (HAP) is more subtle. In fact, it holds for bounded spectrum Ω , where it is a consequence of the Plancherel-Polya inequality for entire functions of exponential growth [48]. However, one can show that (HAP) fails for unbounded spectra, therefore Theorem 2.2 is not directly applicable. In this case one applies the sampling part of the density theorem to the subspace $B_{\Omega \cap B_R(0)} \subseteq B_\Omega$ and then takes the limit $R \rightarrow \infty$. For the interpolation part, we do not need axiom (HAP) - cf. Remark 4.2 - so there is no difficulty for unbounded spectrum Ω , see [34] for the details.

To summarize, Theorem 2.2 implies Landau's fundamental density theorem for bandlimited functions. The geometric properties and all but one property of the kernel are obvious, but the homogeneous approximation property requires some mathematical arguments.

5.2. Functions of Variable Bandwidth. Next we consider the spectral subspaces of the Schrödinger operator $D_q f = -\frac{1}{4\pi^2} f'' + qf$ in dimension 1 with a compactly supported potential $q \in C^2$. Let $\Omega \subseteq \mathbb{R}^+$ be a bounded set and let $PW_\Omega(D_q)$ be the spectral subspace corresponding to spectrum Ω . If $q \equiv 0$, then $D_0 = -\frac{1}{4\pi^2} \frac{d^2}{dx^2}$ is diagonalized by the Fourier transform \mathcal{F} so that $\mathcal{F}D_0\mathcal{F}^{-1}f(\xi) = \xi^2 \hat{f}(\xi)$ is the operator of multiplication by ξ^2 . For the spectral subspace $PW_\Omega(D_0)$ only the spectral values $\xi^2 \in \Omega$ are relevant, therefore $PW_\Omega(D_0) = \{f \in L^2(\mathbb{R}) : \text{supp } \hat{f} \subseteq \Omega^{1/2}\} = B_{\Omega^{1/2}}$ is the Paley-Wiener space of bandlimited functions with spectrum in $\Omega^{1/2} = \{\xi \in \mathbb{R} : \xi^2 \in \Omega\}$. One can show that $PW_\Omega(D_q)$ is a reproducing kernel Hilbert space.

If $q \not\equiv 0$, then we may consider $PW_\Omega(D_q)$ as a perturbation of the Paley-Wiener space. It is therefore natural to expect that the same density conditions for sampling and interpolation also hold for $PW_\Omega(D_q)$. Indeed, in [23] we proved the following result.

Theorem 5.1. *Assume that $\Omega \subseteq \mathbb{R}^+$ is a bounded set with positive (Lebesgue) measure.*

- (i) *If Λ is a set of sampling for $PW_\Omega(D_q)$, then $D^-(\Lambda) \geq \mu(\Omega^{1/2})$.*
- (ii) *If Λ is a set of interpolation for $PW_\Omega(D_q)$, then $D^+(\Lambda) \leq \mu(\Omega^{1/2})$.*

Whereas this result is expected, its proof is surprisingly difficult. In contrast to the Paley-Wiener space $B_\Omega = PW_\Omega(D_0)$, the reproducing kernel for $PW_\Omega(D_q)$ is not known explicitly. To derive Theorem 5.1, we had to use the fine details of the scattering theory for one-dimensional Schrödinger operators for the verification of

the kernel axioms (WL) and (HAP) and for the computation of the averaged trace of the kernel.

Thus for this example our main efforts in [23] were devoted to deriving suitable kernel estimates.

Remark 5.2. (i) For $PW_\Omega(D_q)$ one can also derive sufficient conditions for sampling. See [37] for a qualitative sampling theorem and [23] for an explicit almost optimal sampling theorem.

(ii) In [23] we treated a unitarily equivalent model of the Paley-Wiener space and studied a new concept of variable bandwidth. In that case the density results are formulated differently and also involve a different geometry.

5.3. Sampling in locally compact groups. Let \mathcal{G} be a locally compact group with Haar measure $d\mu = dx$. We make the following additional assumptions:

- (i) \mathcal{G} is compactly generated, i.e., there exists a symmetric neighborhood $U = U^{-1}$ of e with compact closure such that $\mathcal{G} = \bigcup_{n=0}^{\infty} U^n$. The corresponding metric on \mathcal{G} , the so-called word metric, is defined as

$$d(x, y) := \min\{n \in \mathbb{N}_0 : x^{-1}y \in U^n\}, \quad x, y \in \mathcal{G}.$$

It is clearly left-invariant, and the balls are compact sets, in particular Borel sets of finite measure.

- (ii) \mathcal{G} has polynomial growth, i.e., there exist constants $C, D > 0$ such that

$$(39) \quad \mu(U^n) \leq Cn^D, \quad n \in \mathbb{N}.$$

Under these assumptions the word metric possesses the weak annular decay property. In fact, Tessera [46, Cor. 10] showed that polynomial growth implies the annular decay property. Thus (\mathcal{G}, d, μ) satisfies all geometric axioms. (See also [8, 10].)

Now let π be an irreducible, unitary, *square-integrable* representation on a Hilbert space \mathcal{H} . The orthogonality relations for square-integrable representations [17] allow us to identify \mathcal{H} with a reproducing kernel Hilbert space. Precisely, fix a non-zero $g \in \mathcal{H}$ with normalization $\|g\| = 1$ and consider the map

$$\mathcal{C} : \mathcal{H} \rightarrow L^2(\mathcal{G}), \quad \mathcal{C}f(x) = \langle f, \pi(x)g \rangle, \quad x \in \mathcal{G}.$$

The orthogonality relations then imply that

$$(40) \quad \langle \mathcal{C}f, \mathcal{C}h \rangle_{L^2(\mathcal{G})} = d_\pi^{-1} \langle f, h \rangle_{\mathcal{H}},$$

where the constant d_π is the so-called formal dimension of π . Consequently \mathcal{C} is a multiple of an isometry, and we can identify the representation space \mathcal{H} with the subspace $\tilde{\mathcal{H}} = \mathcal{C}\mathcal{H}$ of $L^2(\mathcal{G})$. Now choosing $h = \pi(x)g$ for $x \in \mathcal{G}$ in (40), we obtain that

$$\mathcal{C}f(x) = \langle f, \pi(x)g \rangle_{\mathcal{H}} = d_\pi \langle \mathcal{C}f, \mathcal{C}(\pi(x)g) \rangle_{L^2(\mathcal{G})} = d_\pi \int_{\mathcal{G}} \mathcal{C}f(y) \overline{\langle \pi(x)g, \pi(y)g \rangle} dy.$$

This identity says that $\tilde{\mathcal{H}}$ is a reproducing kernel Hilbert space with kernel

$$k(x, y) = d_\pi \overline{\langle \pi(x)g, \pi(y)g \rangle} = d_\pi \langle g, \pi(y^{-1}x)g \rangle.$$

Consequently, $k(x, x) = d_\pi$ is constant, and axiom D is satisfied trivially. The computation of the averaged trace is a banality and yields

$$\mathrm{tr}^\pm(k) = \frac{1}{\mu(B_r(x))} \int_{B_r(x)} k(y, y) d\mu(y) = d_\pi.$$

Moreover, since $x \rightarrow \langle g, \pi(x)g \rangle$ is in $L^2(\mathcal{G})$, the weak localization (WL) is also satisfied. Again, the homogeneous approximation property (HAP) is the least obvious property and requires some work. Let \mathbf{B} consist of all vectors $g \in \mathcal{H}$ of the form $g = \int_{\mathcal{G}} \eta(x) \pi(x) g_0 d\mu(x)$ for some $g_0 \in \mathcal{H}$ and η a compactly supported continuous function on \mathcal{G} . If $g \in \mathbf{B}$ and $\Lambda \subseteq \mathcal{G}$ is an arbitrary relatively separated set, then the set of reproducing kernels $\{\langle g, \pi(\lambda^{-1}\cdot)g \rangle : \lambda \in \Lambda\}$ satisfies axiom (HAP) by an observation in [22].

To formulate Theorem 2.2 for this particular example, we finally note that $\Lambda \subseteq \mathcal{G}$ is a set of sampling (set of interpolation) for $\tilde{\mathcal{H}}$ if and only if $\{\pi(\lambda)g : \lambda \in \Lambda\}$ is a frame (Riesz sequence) for \mathcal{H} . Frames of this form are often called coherent frames or discrete subsets of coherent states. Theorem 2.2 yields the following density result for coherent frames.

Theorem 5.3. *Let \mathcal{G} be a compactly generated, locally compact group with polynomial growth, and let π be an irreducible, unitary, square-integrable representation on a Hilbert space \mathcal{H} .*

- (i) *If $\{\pi(\lambda)g : \lambda \in \Lambda\}$ is a frame for \mathcal{H} for $g \in \mathbf{B}$, then $D^-(\Lambda) \geq d_\pi$.*
- (ii) *If $\{\pi(\lambda)g : \lambda \in \Lambda\}$ is a Riesz sequence in \mathcal{H} , then $D^+(\Lambda) \leq d_\pi$.*

This result seems to be new. For square-integrable representations of groups of polynomial growth it provides a critical density that separates frames from Riesz sequences. For concrete representations, e.g., the Schrödinger representation of the Heisenberg group Theorem 5.3 has been derived many times in the context of Gabor analysis [25]. For homogeneous (nilpotent) groups it has been proved in the thesis of A. Höfler [26] by using the techniques of Ramanathan and Steger [40].

Let us mention that the construction of coherent frames associated to irreducible representations was first studied systematically in coorbit theory, see [15, 21]. If $\Lambda \subseteq \mathcal{G}$ is “sufficiently dense”, then $\{\pi(\lambda)g : \lambda \in \Lambda\}$ is a frame for \mathcal{H} . Theorem 5.3 complements the existence of such frames by a critical density.

Theorem 2.2 also yields several new density results about sampling and interpolation in reproducing kernel Hilbert spaces that are invariant under a group action. As the full exploitation of Theorem 2.2 is beyond the scope of this section, we will come back to it in further work.

5.4. Complex analysis. Finally, we deal with sampling and interpolation in weighted spaces of analytic functions. We will partially rederive Lindholm's result [31] from Theorem 2.2.

Let ϕ be a plurisubharmonic function on \mathbb{C}^n which is 2-homogeneous and C^2 on $\mathbb{C}^n \setminus \{0\}$. We also assume that there exist $A, B > 0$ such that

$$(41) \quad A \cdot \text{Id}_n \leq \left(\partial_j \bar{\partial}_k \phi(z) \right)_{j,k=1,\dots,n} \leq B \cdot \text{Id}_n$$

for all $z \neq 0$, in the sense of positive definite matrices. It follows that $A^n \leq \det(\partial_j \bar{\partial}_k \phi)_{jk} \leq B^n$ on $\mathbb{C}^n \setminus \{0\}$. Note that in dimension $n = 1$ this condition simply means that the Laplacian $\Delta \phi = \partial \bar{\partial} \phi$ is bounded above and below from 0.

Our main object is the Hilbert space \mathcal{F}_ϕ^2 of entire functions on \mathbb{C}^n defined by the norm $\|f\|_{\mathcal{F}_\phi^2}^2 = \int_{\mathbb{C}^n} |f|^2 e^{-2\phi} dm$. The standard example is the weight $\phi(z) = |z|^2/2$ which yields the Bargmann-Fock space. This is a reproducing kernel Hilbert space with kernel $\frac{1}{\pi} e^{z \cdot \bar{w}}$. Since this kernel is unbounded, we use a different normalization to put it into the framework of Theorem 2.2.

We take $X = \mathbb{C}^n$ with the usual Euclidean distance and the measure $d\mu = \det(\partial_j \bar{\partial}_k \phi)_{jk} dm$ where dm is the Lebesgue measure. Thus μ is equivalent to Lebesgue measure. Let A_ϕ^2 consist of all functions of the form

$$(42) \quad g = \frac{1}{\sqrt{\det(\partial_j \bar{\partial}_k \phi)_{jk}}} f e^{-\phi},$$

where f is entire, such that

$$(43) \quad \|g\|^2 = \int_{\mathbb{C}^n} |g|^2 d\mu = \int_{\mathbb{C}^n} |f|^2 e^{-2\phi} dm < \infty.$$

We observe immediately that the assumptions on the metric and the measure required in the main theorem are satisfied.

It can be shown that A_ϕ^2 is a reproducing kernel Hilbert space. We denote the kernel by $K = K_\phi$. For the weight $\phi(z) = |z|^2/2$, our normalization yields the following explicit expression for the kernel

$$K(z, w) = \frac{2^n}{\pi^n} e^{z \cdot \bar{w} - |z|^2/2 - |w|^2/2}.$$

It is easy to see that this kernel satisfies the axioms (D), (WL), and (HAP), therefore the Seip's necessary density conditions [42] for sampling in Bargmann-Fock space follow (without strict inequalities) directly from Theorem 2.2.

For more general weights ϕ there is no explicit formula for the kernel, but strong estimates are known. We use Lindholm's estimates [31]. Since he works with the measure $e^{-\phi} dm$ and entire functions, we have to translate these results to our normalization with the measure μ and functions of the form (42). Using the observation (20), the relation between the kernel B_ϕ in [31] and our kernel K_ϕ is

given by

$$K_\phi(z, w) = \frac{1}{\sqrt{\det(\partial_j \bar{\partial}_k \phi)_{jk}(z) \det(\partial_j \bar{\partial}_k \phi)_{jk}(w)}} B_\phi(z, w) e^{-\phi(z) - \phi(w)}.$$

Translated to our notation, Lindholm [31] proved the following facts about K_ϕ :

(i) There exist constants $C, T > 0$ depending on A and B in (41), such that for all $k > 0$ holds the decay estimate

$$(44) \quad K_{k^2\phi}(z, w) \leq C e^{-kT|z-w|}.$$

(ii) On the diagonal the kernel satisfies the limit relation

$$(45) \quad \lim_{k \rightarrow \infty} K_{k^2\phi}(z, z) = \frac{2^n}{\pi^n},$$

with uniform convergence on $\mathbb{C}^n \setminus B_\tau(0)$ for arbitrary $\tau > 0$.

In addition, the 2-homogeneity and a simple change of variables imply that

$$(46) \quad K_{k^2\phi}(z, w) = K_\phi(kz, kw).$$

The axioms (WL) and (HAP) follow immediately from the off-diagonal decay (44) of the kernel K_ϕ , likewise $K_\phi(z, z)$ is bounded. The lower bound in axiom (D) can now be verified in the following way. Note first that since $g = (\det(\partial_j \bar{\partial}_k \phi)_{jk})^{-1/2} e^{-\phi} \in A_\phi^2$, it follows that $K(z, z) \neq 0$, for all $z \in \mathbb{C}^n$. In addition, by (45) and (46), there exist $c, \tau, R > 0$ such that $K_{R^2\phi}(z, z) = K(Rz, Rz) > c$ on $\mathbb{C}^n \setminus B_\tau(0)$. Then $K_\phi(z, z) = K_\phi(R \cdot z/R, R \cdot z/R) > c$ for any $|z| > R\tau$. By continuity and the fact that $K(z, z) \neq 0$ for all z , we also have $K(z, z) > c_2$ for some $c_2 > 0$ for $|z| \leq R\tau$.

Thus the geometry and the kernel satisfy all required hypotheses of Section 2.1. Thus Theorem 2.2 is applicable and yields a critical density that separates sampling from interpolation. It remains to compute this critical density.

Lemma 5.4. *If the weight ϕ is plurisubharmonic, 2-homogeneous, and satisfies (41), then*

$$(47) \quad \text{tr}^+(K_\phi) = \text{tr}^-(K_\phi) = \frac{2^n}{\pi^n}.$$

Proof. We use the homogeneity (46) and $(\partial_j \bar{\partial}_k \phi)(rz) = \partial_j \bar{\partial}_k \phi(z)$ for all $r > 0$. Then

$$(48) \quad \begin{aligned} & \sup_{x \in \mathbb{C}^n} \frac{1}{\int_{B_r(x)} \det(\partial_j \bar{\partial}_k \phi)_{jk} dm} \int_{B_r(x)} K_\phi(z, z) d\mu(z) \\ &= \sup_{x \in \mathbb{C}^n} \frac{1}{\int_{B_1(x/r)} \det(\partial_j \bar{\partial}_k \phi)_{jk} dm} \int_{B_1(x/r)} K_{r^2\phi}(w, w) d\mu(w) \\ &= \sup_{y \in \mathbb{C}^n} \frac{1}{\int_{B_1(y)} \det(\partial_j \bar{\partial}_k \phi)_{jk} dm} \int_{B_1(y)} K_{r^2\phi}(w, w) d\mu(w). \end{aligned}$$

Now, we use the fact that $K_{r^2\phi}(w, w)$ converges to $\frac{2^n}{\pi^n}$ uniformly outside any ball $B_\tau(0)$, $\tau > 0$. If $B_1(y)$ contains the origin, we remove a small neighborhood of 0, otherwise we use (45) directly. Given $\epsilon > 0$, it follows that

$$\left| \sup_{x \in \mathbb{C}^n} \frac{1}{\int_{B_r(x)} \det(\partial_j \bar{\partial}_k \phi)_{jk} dm} \int_{B_r(x)} K_{r^2\phi}(z, z) d\mu(z) - \frac{2^n}{\pi^n} \right| \leq \epsilon + o(1)$$

as $r \rightarrow \infty$. Therefore,

$$\limsup_{r \rightarrow \infty} \left| \sup_{x \in \mathbb{C}^n} \frac{1}{\int_{B_r(x)} \det(\partial_j \bar{\partial}_k \phi)_{jk} dm} \int_{B_r(x)} K_\phi(z, z) dm(z) - \frac{2^n}{\pi^n} \right| \leq \epsilon$$

for all $\epsilon > 0$, which means that

$$\text{tr}^+(K_\phi) = \lim_{r \rightarrow \infty} \sup_{x \in \mathbb{C}^n} \frac{1}{\int_{B_r(x)} \det(\partial_j \bar{\partial}_k \phi)_{jk} dm} \int_{B_r(x)} K_\phi(z, z) dm(z) = \frac{2^n}{\pi^n}.$$

Likewise $\text{tr}^-(K_\phi) = \frac{2^n}{\pi^n}$. ■

Theorem 2.2 now implies Lindholm's result [31].

Theorem 5.5. *Assume that ϕ is plurisubharmonic, 2-homogeneous, and satisfies (41).*

- (i) *If $\Lambda \subseteq \mathbb{C}^d$ is a set of sampling for A_ϕ^2 , then $D^-(\Lambda) \geq \frac{2^n}{\pi^n}$.*
- (ii) *If $\Lambda \subseteq \mathbb{C}^d$ is a set of interpolation for A_ϕ^2 , then $D^+(\Lambda) \leq \frac{2^n}{\pi^n}$.*

Remark 5.6. For generalized Fock spaces in one complex variable, Ortega-Cerdà and Seip [35] and Marco, Massaneda, Ortega-Cerdà [33] proved a density theorem for non-homogeneous weights as well. Although Theorem 2.2 applies, we are (not yet) able to recover their explicit result. This would require to derive a version of Lemma 5.4 for non-homogeneous doubling weights.

5.5. Density of Abstract Frames. Finally, we note certain connections with the density theory for abstract frames [4–6].

Let \mathcal{H} be a separable Hilbert space, (X, d) a countable metric space with counting measure μ , and $\mathcal{F} = \{f_x : x \in X\}$ a frame for \mathcal{H} , i.e., there exist $A, B > 0$ such that

$$(49) \quad A\|f\|^2 \leq \sum_{x \in X} |\langle f, f_x \rangle|^2 \leq B\|f\|^2 \quad \forall f \in \mathcal{H}.$$

Using the coefficient operator $\mathcal{C} : \mathcal{H} \rightarrow \ell^2(X)$

$$(50) \quad \mathcal{C}f(x) = \langle f, f_x \rangle \quad x \in X,$$

we can identify the abstract Hilbert space \mathcal{H} with the subspace of functions $\mathcal{C}f$ of $\ell^2(X)$. By the frame inequalities (49) \mathcal{C} is one-to-one with closed range in $\ell^2(X)$, which we call $\tilde{\mathcal{H}} = \mathcal{C}\mathcal{H} \subseteq \ell^2(X)$.

Let $\{\tilde{f}_x : x \in X\}$ be the (canonical) dual frame of \mathcal{H} , then every $f \in \mathcal{H}$ possesses the frame expansion $f = \sum_{y \in X} \langle f, f_y \rangle \tilde{f}_y$, and consequently

$$\mathcal{C}f(x) = \sum_{y \in X} \langle f, f_y \rangle \langle \tilde{f}_y, f_x \rangle = \sum_{y \in X} \mathcal{C}f(y) \langle \tilde{f}_y, f_x \rangle,$$

This means that $\tilde{\mathcal{H}}$ is a reproducing kernel subspace of $\ell^2(X)$ with kernel

$$k(x, y) = \langle \tilde{f}_y, f_x \rangle.$$

The two properties (WL) and (HAP) for the pair $(X, \tilde{\mathcal{H}})$ are equivalent to what in [4] is called ℓ^2 -localization of the frames \mathcal{F} and $\tilde{\mathcal{F}}$. Furthermore, the (lower) averaged trace of this kernel is

$$\mathrm{tr}^-(k) = \liminf_{r \rightarrow \infty} \inf_{x \in X} \frac{1}{\#B_r(x)} \sum_{y \in B_r(x)} \langle \tilde{f}_y, f_y \rangle.$$

This quantity correspond exactly to the (lower) *frame measure* of \mathcal{F} in [4, 6].

Besides these technical similarities, Theorem 2.2 is not formally comparable to the results in [4–6]. The theory of Balan, Casazza, Heil, and Landau in [4–6] compares two abstract frames, and derives an equality relating density and measure. By contrast, Theorem 2.2 compares a frame of reproducing kernels to a possibly continuous resolution of the identity.

REFERENCES

- [1] L. D. Abreu. Sampling and interpolation in Bargmann-Fock spaces of polyanalytic functions. *Appl. Comput. Harmon. Anal.*, 29(3):287–302, 2010.
- [2] L. D. Abreu and A. S. Bandeira. Landau’s necessary density conditions for the Hankel transform. *J. Funct. Anal.*, 262(4):1845–1866, 2012.
- [3] A. Aldroubi and K. Gröchenig. Nonuniform sampling and reconstruction in shift-invariant spaces. *SIAM Rev.*, 43(4):585–620 (electronic), 2001.
- [4] R. Balan, P. G. Casazza, C. Heil, and Z. Landau. Density, overcompleteness, and localization of frames. I. Theory. *J. Fourier Anal. Appl.*, 12(2):105–143, 2006.
- [5] R. Balan, P. G. Casazza, C. Heil, and Z. Landau. Density, overcompleteness, and localization of frames. II. Gabor systems. *J. Fourier Anal. Appl.*, 12(3):309–344, 2006.
- [6] R. Balan and Z. Landau. Measure functions for frames. *J. Funct. Anal.*, 252(2):630–676, 2007.
- [7] A. Beurling. Local harmonic analysis with some applications to differential operators. In *Some Recent Advances in the Basic Sciences, Vol. 1 (Proc. Annual Sci. Conf., Belfer Grad. School Sci., Yeshiva Univ., New York, 1962–1964)*, pages 109–125. Belfer Graduate School of Science, Yeshiva Univ., New York, 1966.
- [8] E. Breuillard. Geometry of locally compact groups of polynomial growth and shape of large balls. *Groups Geom. Dyn.*, 8(3):669–732, 2014.
- [9] S. M. Buckley. Is the maximal function of a Lipschitz function continuous? *Ann. Acad. Sci. Fenn. Math.*, 24(2):519–528, 1999.
- [10] T. H. Colding and W. P. Minicozzi, II. Liouville theorems for harmonic sections and applications. *Comm. Pure Appl. Math.*, 51(2):113–138, 1998.

- [11] T. Coulhon and L. Saloff-Coste. Variétés riemanniennes isométriques à l'infini. *Rev. Mat. Iberoamericana*, 11(3):687–726, 1995.
- [12] I. Daubechies. The wavelet transform, time-frequency localization and signal analysis. *IEEE Trans. Inform. Theory*, 36(5):961–1005, 1990.
- [13] R. J. Duffin and A. C. Schaeffer. A class of nonharmonic Fourier series. *Trans. Amer. Math. Soc.*, 72:341–366, 1952.
- [14] H. G. Feichtinger and K. Gröchenig. Theory and practice of irregular sampling. In *Wavelets: mathematics and applications*, Stud. Adv. Math., pages 305–363. CRC, Boca Raton, FL, 1994.
- [15] H. G. Feichtinger and K. H. Gröchenig. Banach spaces related to integrable group representations and their atomic decompositions. I. *J. Funct. Anal.*, 86(2):307–340, 1989.
- [16] H. G. Feichtinger, H. Führ, I. Z. Pesenson. Geometric space-frequency analysis on manifolds. Preprint. arXiv:1512.08668.
- [17] H. Führ. *Abstract harmonic analysis of continuous wavelet transforms*, volume 1863 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2005.
- [18] H. Führ. Simultaneous estimates for vector-valued Gabor frames of Hermite functions. *Adv. Comput. Math.*, 29(4):357–373, 2008.
- [19] H. Führ and K. Gröchenig. Sampling theorems on locally compact groups from oscillation estimates. *Math. Z.*, 255(1):177–194, 2007.
- [20] C. C. Graham. A uniform boundedness principle for compact sets and the decay of Fourier transforms. *Math. Z.*, 196(3):331–342, 1987.
- [21] K. Gröchenig. Describing functions: atomic decompositions versus frames. *Monatsh. Math.*, 112(1):1–42, 1991.
- [22] K. Gröchenig. The homogeneous approximation property and the comparison theorem for coherent frames. *Sampl. Theory Signal Image Process.*, 7(3):271–279, 2008.
- [23] K. Gröchenig and A. Klotz. What is Variable Bandwidth? *ArXiv e-prints*, arXiv 1512.06663v1, Dec. 2015.
- [24] K. Gröchenig and H. Razafinjato. On Landau's necessary density conditions for sampling and interpolation of band-limited functions. *J. London Math. Soc. (2)*, 54(3):557–565, 1996.
- [25] C. Heil. History and evolution of the density theorem for Gabor frames. *J. Fourier Anal. Appl.*, 13(2):113–166, 2007.
- [26] A. Höfler. *Necessary density conditions for frames on homogeneous groups*. PhD thesis, Universität Wien, 2014.
- [27] A. Iosevich and M. N. Kolountzakis. A Weyl type formula for Fourier spectra and frames. *Proc. Amer. Math. Soc.*, 134(11):3267–3274 (electronic), 2006.
- [28] M. N. Kolountzakis and J. C. Lagarias. Structure of tilings of the line by a function. *Duke Math. J.*, 82(3):653–678, 1996.
- [29] G. Kutyniok. *Affine density in wavelet analysis*, volume 1914 of *Lecture Notes in Mathematics*. Springer, Berlin, 2007.
- [30] H. J. Landau. Necessary density conditions for sampling and interpolation of certain entire functions. *Acta Math.*, 117:37–52, 1967.
- [31] N. Lindholm. Sampling in weighted L^p spaces of entire functions in \mathbb{C}^n and estimates of the Bergman kernel. *J. Funct. Anal.*, 182(2):390–426, 2001.
- [32] Y. I. Lyubarskiĭ. Frames in the Bargmann space of entire functions. In *Entire and subharmonic functions*, pages 167–180. Amer. Math. Soc., Providence, RI, 1992.
- [33] N. Marco, X. Massaneda, and J. Ortega-Cerdà. Interpolating and sampling sequences for entire functions. *Geometric & Functional Analysis GAFA*, 13(4):862–914, 2003.

- [34] S. Nitzan and A. Olevskii. Revisiting Landau's density theorems for Paley-Wiener spaces. *C. R. Math. Acad. Sci. Paris*, 350(9-10):509–512, 2012.
- [35] J. Ortega-Cerdà and K. Seip. Beurling-type density theorems for weighted l^p spaces of entire functions. *J. Anal. Math.*, 75(1):247–266, 1998.
- [36] I. Pesenson. Sampling of Paley-Wiener functions on stratified groups. *J. Fourier Anal. Appl.*, 4(3):271–281, 1998.
- [37] I. Pesenson. Sampling of band-limited vectors. *J. Fourier Anal. Appl.*, 7(1):93–100, 2001.
- [38] I. Pesenson and A. I. Zayed. Paley-Wiener subspace of vectors in a Hilbert space with applications to integral transforms. *J. Math. Anal. Appl.*, 353(2):566–582, 2009.
- [39] G. Kerkycharian and P. Petrushev. Heat kernel based decomposition of spaces of distributions in the framework of Dirichlet spaces. *Trans. Amer. Math. Soc.*, 367(1):121–189, 2015.
- [40] J. Ramanathan and T. Steger. Incompleteness of sparse coherent states. *Appl. Comput. Harmon. Anal.*, 2(2):148–153, 1995.
- [41] A. P. Schuster. On Seip's description of sampling sequences for Bergman spaces. *Complex Variables Theory Appl.*, 42(4):347–367, 2000.
- [42] K. Seip. Density theorems for sampling and interpolation in the Bargmann-Fock space. I. *J. Reine Angew. Math.*, 429:91–106, 1992.
- [43] K. Seip. Beurling type density theorems in the unit disk. *Invent. Math.*, 113(1):21–39, 1993.
- [44] K. Seip. *Interpolation and sampling in spaces of analytic functions*, volume 33 of *University Lecture Series*. American Mathematical Society, Providence, RI, 2004.
- [45] K. Seip and R. Wallstén. Density theorems for sampling and interpolation in the Bargmann-Fock space. II. *J. Reine Angew. Math.*, 429:107–113, 1992.
- [46] R. Tessera. Volume of spheres in doubling metric measured spaces and in groups of polynomial growth. *Bull. Soc. Math. France*, 135(1):47–64, 2007.
- [47] R. Tessera. Large scale Sobolev inequalities on metric measure spaces and applications. *Rev. Mat. Iberoam.*, 24(3):825–864, 2008.
- [48] H. Triebel. *Theory of function spaces*. Birkhäuser Verlag, Basel, 1983.
- [49] M. Unser. Sampling - 50 Years after Shannon. *Proc. IEEE*, 88 (4):569–587, 2000.

FACULTY OF MATHEMATICS, UNIVERSITY OF VIENNA, OSKAR-MORGENSTERN-PLATZ 1,
A-1090 VIENNA, AUSTRIA

E-mail address: fuehr@matha.rwth-aachen.de

E-mail address: karlheinz.groechnig@univie.ac.at

E-mail address: antti.haimi@univie.ac.at

E-mail address: andreas.klotz@univie.ac.at

E-mail address: jose.luis.romero@univie.ac.at